

TRhizo-urbanMicrobiome

Ecosystem Function Bacteria

David Murray-Stoker

Contents

| | |
|-----------------------------------|-----------|
| Load Packages | 3 |
| Load Data | 4 |
| Data Management | 5 |
| Filter Abundances | 6 |
| Nitrosomonas GAMs | 10 |
| Check Model Assumptions | 10 |
| ANOVAs | 11 |
| Nitrospira GAMs | 12 |
| Check Model Assumptions | 12 |
| ANOVAs | 13 |
| Desulfuromonas GAMs | 14 |
| Check Model Assumptions | 14 |
| ANOVAs | 15 |
| Aciditerrimonas GAMs | 16 |
| Check Model Assumptions | 16 |
| ANOVAs | 17 |
| Nitrosospira GAMs | 18 |
| Check Model Assumptions | 18 |
| ANOVAs | 19 |
| Sphingopyxis GAMs | 20 |
| Check Model Assumptions | 20 |
| ANOVAs | 21 |
| Cytophaga GAMs | 22 |
| Check Model Assumptions | 22 |
| ANOVAs | 23 |
| Paracoccus GAMs | 24 |
| Check Model Assumptions | 24 |
| ANOVAs | 25 |
| Novosphingobium GAMs | 26 |
| Check Model Assumptions | 26 |

| | |
|-----------------------------------|-----------|
| ANOVAs | 27 |
| Devosia GAMs | 28 |
| Check Model Assumptions | 28 |
| ANOVAs | 29 |
| Streptomyces GAMs | 30 |
| Check Model Assumptions | 30 |
| ANOVAs | 31 |
| Sphingobium GAMs | 32 |
| Check Model Assumptions | 32 |
| ANOVAs | 33 |
| Desulfosporosinus GAMs | 34 |
| Check Model Assumptions | 34 |
| ANOVAs | 35 |
| Desulfocapsa GAMs | 36 |
| Check Model Assumptions | 36 |
| ANOVAs | 37 |
| Phyllobacterium GAMs | 38 |
| Check Model Assumptions | 38 |
| ANOVAs | 39 |
| Export Data | 40 |
| R Session Information | 41 |

Load Packages

Load Data

```
## Microbiome tidyamplicons
microbiome.tidyamplicon.reference <- read_rds(
  file = "data/microbiome_tidyamplicon_reference.rds"
)

## Load the urbanization data
urbanization.data <- read_rds(
  file = "data/urbanization_data.rds"
)

## Load the workspace
#load("data_analysis/10-bacteria_functional_groups/ecosystem_function_bacteria-workspace.RData")
```

Data Management

```
## Set reference ASV abundance matrix
ASV.abundance.matrix <- add_rel_abundance(microbiome.tidyamplicon.reference) %>%
  abundances()

## Set reference taxa matrix
ASV.taxa.matrix <- taxa(microbiome.tidyamplicon.reference)

## Set tibble of bacteria abundance
bacteria.abundance.data <- ASV.abundance.matrix %>%
  full_join(ASV.taxa.matrix, by = "taxon_id") %>%
  full_join(microbiome.tidyamplicon.reference$samples, by = "sample_id") %>%
  select(Population:Compartment, phylum:genus, abundance, rel_abundance) %>%
  rename(
    Phylum = phylum, Class = class, Order = order, Family = family, Genus = genus,
    Abundance = abundance, Relative_Abundance = rel_abundance
  )

## Set aggregated sample data
sample.data <- microbiome.tidyamplicon.reference$samples %>%
  select(Population:Compartment) %>%
  distinct()
```

Filter Abundances

```
## Set tibbles of abundances for different mutualistic bacteria
# Nitrosomonas
nitrosomonas.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Nitrosomonas") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Nitrosomonas_Abundance = sum(Abundance),
    Nitrosomonas_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Nitrospira
nitrospira.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Nitrospira") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Nitrospira_Abundance = sum(Abundance),
    Nitrospira_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Desulfuromonas
desulfuromonas.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Desulfuromonas") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Desulfuromonas_Abundance = sum(Abundance),
    Desulfuromonas_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Aciditerrimonas
aciditerrimonas.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Aciditerrimonas") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Aciditerrimonas_Abundance = sum(Abundance),
    Aciditerrimonas_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Nitrosospira
nitrosospira.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Nitrosospira") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
```

```

    Nitrosospira_Abundance = sum(Abundance),
    Nitrosospira_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
)

# Sphingopyxis
sphingopyxis.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Sphingopyxis") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Sphingopyxis_Abundance = sum(Abundance),
    Sphingopyxis_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
)

# Cytophaga
cytophaga.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Cytophaga") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Cytophaga_Abundance = sum(Abundance),
    Cytophaga_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
)

# Paracoccus
paracoccus.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Paracoccus") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Paracoccus_Abundance = sum(Abundance),
    Paracoccus_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
)

# Novosphingobium
novosphingobium.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Novosphingobium") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Novosphingobium_Abundance = sum(Abundance),
    Novosphingobium_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
)

# Devosia
devosia.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Devosia") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%

```

```

group_by(Population, Compartment) %>%
  summarise(
    Devosia_Abundance = sum(Abundance),
    Devosia_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Streptomyces
streptomyces.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Streptomyces") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Streptomyces_Abundance = sum(Abundance),
    Streptomyces_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Sphingobium
sphingobium.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Sphingobium") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Sphingobium_Abundance = sum(Abundance),
    Sphingobium_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Desulfosporosinus
desulfosporosinus.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Desulfosporosinus") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Desulfosporosinus_Abundance = sum(Abundance),
    Desulfosporosinus_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Desulfocapsa
desulfocapsa.abundances <- bacteria.abundance.data %>%
  filter(Genus == "Desulfocapsa") %>%
  select(Population, Compartment, Abundance, Relative_Abundance) %>%
  group_by(Population, Compartment) %>%
  summarise(
    Desulfocapsa_Abundance = sum(Abundance),
    Desulfocapsa_Relative_Abundance = sum(Relative_Abundance),
    .groups = "keep"
  )

# Phyllobacterium
phyllobacterium.abundances <- bacteria.abundance.data %>%

```


Nitrosomonas GAMs

```
## Nitrosomonas by distance
nitrosomonas.by.distance.GAM <- gam(
  Nitrosomonas_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrosomonas by HII
nitrosomonas.by.HII.GAM <- gam(
  Nitrosomonas_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrosomonas by ISC
nitrosomonas.by.ISC.GAM <- gam(
  Nitrosomonas_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Nitrosomonas abundance-by-distance model diagnostics
check_model(nitrosomonas.by.distance.GAM)
# Visual check = assumptions met

## Nitrosomonas abundance-by-HII model diagnostics
check_model(nitrosomonas.by.HII.GAM)
# Visual check = assumptions met

## Nitrosomonas abundance-by-ISC model diagnostics
check_model(nitrosomonas.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 1: ANOVA table for the Nitrosomonas relative abundance-by-distance GAM. Adjusted R-squared = 0.530, deviance = 55.5
Compartment: F = 109.3, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.531 |
| s(Distance):CompartmentSoil | 2.644 | 9 | 7.068 | 0.000 |

Table 2: ANOVA table for the Nitrosomonas relative abundance-by-HII GAM. Adjusted R-squared = 0.193, deviance = 21.5
Compartment: F = 15.29, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.821 |
| s(Human_Influence_Index):CompartmentSoil | 0.896 | 9 | 0.952 | 0.003 |

Table 3: ANOVA table for the Nitrosomonas relative abundance-by-ISC GAM. Adjusted R-squared = 0.219, deviance = 24.1
Compartment: F = 8.895, P = 0.004.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.00 | 9 | 0.000 | 0.860 |
| s(Mean_ISC):CompartmentSoil | 0.96 | 9 | 1.236 | 0.001 |

Nitrospira GAMs

```
## Nitrospira by distance
nitrospira.by.distance.GAM <- gam(
  Nitrospira_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrospira by HII
nitrospira.by.HII.GAM <- gam(
  Nitrospira_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrospira by ISC
nitrospira.by.ISC.GAM <- gam(
  Nitrospira_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Nitrospira abundance-by-distance model diagnostics
check_model(nitrospira.by.distance.GAM)
# Visual check = assumptions met

## Nitrospira abundance-by-HII model diagnostics
check_model(nitrospira.by.HII.GAM)
# Visual check = assumptions met

## Nitrospira abundance-by-ISC model diagnostics
check_model(nitrospira.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 4: ANOVA table for the Nitrospira relative abundance-by-distance GAM. Adjusted R-squared = 0.861, deviance = 86.8 Compartment: $F = 354.2$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.789 |
| s(Distance):CompartmentSoil | 2.403 | 9 | 9.207 | 0.000 |

Table 5: ANOVA table for the Nitrospira relative abundance-by-HII GAM. Adjusted R-squared = 0.808, deviance = 82.0 Compartment: $F = 250.7$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.901 |
| s(Human_Influence_Index):CompartmentSoil | 3.051 | 9 | 4.616 | 0.000 |

Table 6: ANOVA table for the Nitrospira relative abundance-by-ISC GAM. Adjusted R-squared = 0.762, deviance = 76.9 Compartment: $F = 201.7$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.00 | 0.868 |
| s(Mean_ISC):CompartmentSoil | 1.018 | 9 | 2.24 | 0.000 |

Desulfuromonas GAMs

```
## Desulfuromonas by distance
desulfuromonas.by.distance.GAM <- gam(
  Desulfuromonas_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfuromonas by HII
desulfuromonas.by.HII.GAM <- gam(
  Desulfuromonas_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfuromonas by ISC
desulfuromonas.by.ISC.GAM <- gam(
  Desulfuromonas_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Desulfuromonas abundance-by-distance model diagnostics
check_model(desulfuromonas.by.distance.GAM)
# Visual check = assumptions met

## Desulfuromonas abundance-by-HII model diagnostics
check_model(desulfuromonas.by.HII.GAM)
# Visual check = assumptions met

## Desulfuromonas abundance-by-ISC model diagnostics
check_model(desulfuromonas.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 7: ANOVA table for the *Desulfuromonas* relative abundance-by-distance GAM. Adjusted R-squared = 0.461, deviance = 47.7
Compartment: F = 38.81, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.892 |
| s(Distance):CompartmentSoil | 1.041 | 9 | 2.346 | 0.000 |

Table 8: ANOVA table for the *Desulfuromonas* relative abundance-by-HII GAM. Adjusted R-squared = 0.395, deviance = 41.2
Compartment: F = 34.63, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.00 | 9 | 0.000 | 0.909 |
| s(Human_Influence_Index):CompartmentSoil | 0.92 | 9 | 1.277 | 0.001 |

Table 9: ANOVA table for the *Desulfuromonas* relative abundance-by-ISC GAM. Adjusted R-squared = 0.424, deviance = 44.1
Compartment: F = 36.34, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.000 | 0.92 |
| s(Mean_ISC):CompartmentSoil | 1.003 | 9 | 1.715 | 0.00 |

Aciditerrimonas GAMs

```
## Aciditerrimonas by distance
aciditerrimonas.by.distance.GAM <- gam(
  Aciditerrimonas_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Aciditerrimonas by HII
aciditerrimonas.by.HII.GAM <- gam(
  Aciditerrimonas_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Aciditerrimonas by ISC
aciditerrimonas.by.ISC.GAM <- gam(
  Aciditerrimonas_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Aciditerrimonas abundance-by-distance model diagnostics
check_model(aciditerrimonas.by.distance.GAM)
# Visual check = assumptions met

## Aciditerrimonas abundance-by-HII model diagnostics
check_model(aciditerrimonas.by.HII.GAM)
# Visual check = assumptions met

## Aciditerrimonas abundance-by-ISC model diagnostics
check_model(aciditerrimonas.by.ISC.GAM)
# Visual check = assumptions met
```


ANOVAs

Table 10: ANOVA table for the Aciditerrimonas relative abundance-by-distance GAM. Adjusted R-squared = 0.531, deviance = 54.6 Compartment: F = 42.19, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.989 |
| s(Distance):CompartmentSoil | 1.148 | 9 | 4.109 | 0.000 |

Table 11: ANOVA table for the Aciditerrimonas relative abundance-by-HII GAM. Adjusted R-squared = 0.393, deviance = 41.0 Compartment: F = 32.57, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.00 | 0.885 |
| s(Human_Influence_Index):CompartmentSoil | 0.929 | 9 | 1.45 | 0.000 |

Table 12: ANOVA table for the Aciditerrimonas relative abundance-by-ISC GAM. Adjusted R-squared = 0.357, deviance = 37.5 Compartment: F = 30.77, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.00 | 0.860 |
| s(Mean_ISC):CompartmentSoil | 0.931 | 9 | 0.95 | 0.003 |

Nitrosospira GAMs

```
## Nitrosospira by distance
nitrosospira.by.distance.GAM <- gam(
  Nitrosospira_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrosospira by HII
nitrosospira.by.HII.GAM <- gam(
  Nitrosospira_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Nitrosospira by ISC
nitrosospira.by.ISC.GAM <- gam(
  Nitrosospira_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Nitrosospira abundance-by-distance model diagnostics
check_model(nitrosospira.by.distance.GAM)
# Visual check = assumptions met

## Nitrosospira abundance-by-HII model diagnostics
check_model(nitrosospira.by.HII.GAM)
# Visual check = assumptions met

## Nitrosospira abundance-by-ISC model diagnostics
check_model(nitrosospira.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 13: ANOVA table for the Nitrosospira relative abundance-by-distance GAM. Adjusted R-squared = 0.794, deviance = 80.8 Compartment: $F = 237.1$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.728 |
| s(Distance):CompartmentSoil | 3.494 | 9 | 3.362 | 0.000 |

Table 14: ANOVA table for the Nitrosospira relative abundance-by-HII GAM. Adjusted R-squared = 0.715, deviance = 72.3 Compartment: $F = 171.3$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.910 |
| s(Human_Influence_Index):CompartmentSoil | 0.751 | 9 | 0.334 | 0.049 |

Table 15: ANOVA table for the Nitrosospira relative abundance-by-ISC GAM. Adjusted R-squared = 0.758, deviance = 76.5 Compartment: $F = 201.7$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.00 | 0.791 |
| s(Mean_ISC):CompartmentSoil | 1.002 | 9 | 1.73 | 0.000 |

Sphingopyxis GAMs

```
## Sphingopyxis by distance
sphingopyxis.by.distance.GAM <- gam(
  Sphingopyxis_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Sphingopyxis by HII
sphingopyxis.by.HII.GAM <- gam(
  Sphingopyxis_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Sphingopyxis by ISC
sphingopyxis.by.ISC.GAM <- gam(
  Sphingopyxis_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Sphingopyxis abundance-by-distance model diagnostics
check_model(sphingopyxis.by.distance.GAM)
# Visual check = assumptions met

## Sphingopyxis abundance-by-HII model diagnostics
check_model(sphingopyxis.by.HII.GAM)
# Visual check = assumptions met

## Sphingopyxis abundance-by-ISC model diagnostics
check_model(sphingopyxis.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 16: ANOVA table for the Sphingopyxis relative abundance-by-distance GAM. Adjusted R-squared = 0.551, deviance = 56.1 Compartment: F = 84.73, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.444 | 9 | 0.087 | 0.187 |
| s(Distance):CompartmentSoil | 0.000 | 9 | 0.000 | 0.789 |

Table 17: ANOVA table for the Sphingopyxis relative abundance-by-HII GAM. Adjusted R-squared = 0.545, deviance = 55.2 Compartment: F = 83.76, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-----|---------|---|---------|
| s(Human_Influence_Index):CompartmentRoot | 0 | 9 | 0 | 0.632 |
| s(Human_Influence_Index):CompartmentSoil | 0 | 9 | 0 | 0.708 |

Table 18: ANOVA table for the Sphingopyxis relative abundance-by-ISC GAM. Adjusted R-squared = 0.563, deviance = 57.4 Compartment: F = 87.22, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.75 | 9 | 0.311 | 0.056 |
| s(Mean_ISC):CompartmentSoil | 0.00 | 9 | 0.000 | 0.833 |

Cytophaga GAMs

```
## Cytophaga by distance
cytophaga.by.distance.GAM <- gam(
  Cytophaga_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Cytophaga by HII
cytophaga.by.HII.GAM <- gam(
  Cytophaga_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Cytophaga by ISC
cytophaga.by.ISC.GAM <- gam(
  Cytophaga_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Cytophaga abundance-by-distance model diagnostics
check_model(cytophaga.by.distance.GAM)
# Visual check = assumptions met

## Cytophaga abundance-by-HII model diagnostics
check_model(cytophaga.by.HII.GAM)
# Visual check = assumptions met

## Cytophaga abundance-by-ISC model diagnostics
check_model(cytophaga.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 19: ANOVA table for the Cytophaga relative abundance-by-distance GAM. Adjusted R-squared = 0.554, deviance = 59.9
Compartment: F = 34.08, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.000 | 9 | 0.000 | 0.912 |
| s(Distance):CompartmentSoil | 6.008 | 9 | 5.849 | 0.000 |

Table 20: ANOVA table for the Cytophaga relative abundance-by-HII GAM. Adjusted R-squared = 0.404, deviance = 42.1
Compartment: F = 25.52, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.767 |
| s(Human_Influence_Index):CompartmentSoil | 0.958 | 9 | 2.483 | 0.000 |

Table 21: ANOVA table for the Cytophaga relative abundance-by-ISC GAM. Adjusted R-squared = 0.489, deviance = 51.5
Compartment: F = 29.75, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.000 | 0.901 |
| s(Mean_ISC):CompartmentSoil | 2.495 | 9 | 4.148 | 0.000 |

Paracoccus GAMs

```
## Paracoccus by distance
paracoccus.by.distance.GAM <- gam(
  Paracoccus_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Paracoccus by HII
paracoccus.by.HII.GAM <- gam(
  Paracoccus_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Paracoccus by ISC
paracoccus.by.ISC.GAM <- gam(
  Paracoccus_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Paracoccus abundance-by-distance model diagnostics
check_model(paracoccus.by.distance.GAM)
# Visual check = assumptions met

## Paracoccus abundance-by-HII model diagnostics
check_model(paracoccus.by.HII.GAM)
# Visual check = assumptions met

## Paracoccus abundance-by-ISC model diagnostics
check_model(paracoccus.by.ISC.GAM)
# Visual check = assumptions met
```


ANOVAs

Table 22: ANOVA table for the Paracoccus relative abundance-by-distance GAM. Adjusted R-squared = 0.181, deviance = 20.4
Compartment: F = 6.634, P = 0.012.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.958 | 9 | 1.064 | 0.002 |
| s(Distance):CompartmentSoil | 0.001 | 9 | 0.000 | 0.392 |

Table 23: ANOVA table for the Paracoccus relative abundance-by-HII GAM. Adjusted R-squared = 0.129, deviance = 15.2
Compartment: F = 6.24, P = 0.015.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.832 | 9 | 0.552 | 0.017 |
| s(Human_Influence_Index):CompartmentSoil | 0.001 | 9 | 0.000 | 0.464 |

Table 24: ANOVA table for the Paracoccus relative abundance-by-ISC GAM. Adjusted R-squared = 0.245, deviance = 26.9
Compartment: F = 7.202, P = 0.009.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 1.021 | 9 | 1.783 | 0.000 |
| s(Mean_ISC):CompartmentSoil | 0.152 | 9 | 0.020 | 0.282 |

Novosphingobium GAMs

```
## Novosphingobium by distance
novosphingobium.by.distance.GAM <- gam(
  Novosphingobium_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Novosphingobium by HII
novosphingobium.by.HII.GAM <- gam(
  Novosphingobium_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Novosphingobium by ISC
novosphingobium.by.ISC.GAM <- gam(
  Novosphingobium_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Novosphingobium abundance-by-distance model diagnostics
check_model(novosphingobium.by.distance.GAM)
# Visual check = assumptions met

## Novosphingobium abundance-by-HII model diagnostics
check_model(novosphingobium.by.HII.GAM)
# Visual check = assumptions met

## Novosphingobium abundance-by-ISC model diagnostics
check_model(novosphingobium.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 25: ANOVA table for the Novosphingobium relative abundance-by-distance GAM. Adjusted R-squared = 0.812, deviance = 81.7 Compartment: F = 296.5, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.737 | 9 | 0.279 | 0.068 |
| s(Distance):CompartmentSoil | 0.000 | 9 | 0.000 | 0.818 |

Table 26: ANOVA table for the Novosphingobium relative abundance-by-HII GAM. Adjusted R-squared = 0.805, deviance = 80.8 Compartment: F = 285.9, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-----|---------|---|---------|
| s(Human_Influence_Index):CompartmentRoot | 0 | 9 | 0 | 0.966 |
| s(Human_Influence_Index):CompartmentSoil | 0 | 9 | 0 | 0.784 |

Table 27: ANOVA table for the Novosphingobium relative abundance-by-ISC GAM. Adjusted R-squared = 0.806, deviance = 80.9 Compartment: F = 287.3, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.246 | 9 | 0.036 | 0.254 |
| s(Mean_ISC):CompartmentSoil | 0.000 | 9 | 0.000 | 0.829 |

Devosia GAMs

```
## Devosia by distance
devosia.by.distance.GAM <- gam(
  Devosia_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Devosia by HII
devosia.by.HII.GAM <- gam(
  Devosia_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Devosia by ISC
devosia.by.ISC.GAM <- gam(
  Devosia_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Devosia abundance-by-distance model diagnostics
check_model(devosia.by.distance.GAM)
# Visual check = assumptions met

## Devosia abundance-by-HII model diagnostics
check_model(devosia.by.HII.GAM)
# Visual check = assumptions met

## Devosia abundance-by-ISC model diagnostics
check_model(devosia.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 28: ANOVA table for the Devosia relative abundance-by-distance GAM. Adjusted R-squared = 0.849, deviance = 85.1 Compartment: F = 388.3, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-----|---------|---|---------|
| s(Distance):CompartmentRoot | 0 | 9 | 0 | 0.903 |
| s(Distance):CompartmentSoil | 0 | 9 | 0 | 0.662 |

Table 29: ANOVA table for the Devosia relative abundance-by-HII GAM. Adjusted R-squared = 0.850, deviance = 85.3 Compartment: F = 391.5, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.361 | 9 | 0.063 | 0.215 |
| s(Human_Influence_Index):CompartmentSoil | 0.000 | 9 | 0.000 | 0.656 |

Table 30: ANOVA table for the Devosia relative abundance-by-ISC GAM. Adjusted R-squared = 0.849, deviance = 85.1 Compartment: F = 388.3, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-----|---------|---|---------|
| s(Mean_ISC):CompartmentRoot | 0 | 9 | 0 | 0.483 |
| s(Mean_ISC):CompartmentSoil | 0 | 9 | 0 | 0.818 |

Streptomyces GAMs

```
## Streptomyces by distance
streptomyces.by.distance.GAM <- gam(
  Streptomyces_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Streptomyces by HII
streptomyces.by.HII.GAM <- gam(
  Streptomyces_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Streptomyces by ISC
streptomyces.by.ISC.GAM <- gam(
  Streptomyces_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Streptomyces abundance-by-distance model diagnostics
check_model(streptomyces.by.distance.GAM)
# Visual check = assumptions met

## Streptomyces abundance-by-HII model diagnostics
check_model(streptomyces.by.HII.GAM)
# Visual check = assumptions met

## Streptomyces abundance-by-ISC model diagnostics
check_model(streptomyces.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 31: ANOVA table for the Streptomyces relative abundance-by-distance GAM. Adjusted R-squared = 0.533, deviance = 54.5 Compartment: $F = 78.17$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.646 | 9 | 0.192 | 0.105 |
| s(Distance):CompartmentSoil | 0.000 | 9 | 0.000 | 0.860 |

Table 32: ANOVA table for the Streptomyces relative abundance-by-HII GAM. Adjusted R-squared = 0.522, deviance = 52.9 Compartment: $F = 76.23$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|--|-----|---------|---|---------|
| s(Human_Influence_Index):CompartmentRoot | 0 | 9 | 0 | 0.694 |
| s(Human_Influence_Index):CompartmentSoil | 0 | 9 | 0 | 0.938 |

Table 33: ANOVA table for the Streptomyces relative abundance-by-ISC GAM. Adjusted R-squared = 0.522, deviance = 52.9 Compartment: $F = 76.23$, $P < 0.001$.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-----|---------|---|---------|
| s(Mean_ISC):CompartmentRoot | 0 | 9 | 0 | 0.822 |
| s(Mean_ISC):CompartmentSoil | 0 | 9 | 0 | 0.997 |

Sphingobium GAMs

```
## Sphingobium by distance
sphingobium.by.distance.GAM <- gam(
  Sphingobium_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Sphingobium by HII
sphingobium.by.HII.GAM <- gam(
  Sphingobium_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Sphingobium by ISC
sphingobium.by.ISC.GAM <- gam(
  Sphingobium_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Sphingobium abundance-by-distance model diagnostics
check_model(sphingobium.by.distance.GAM)
# Visual check = assumptions met

## Sphingobium abundance-by-HII model diagnostics
check_model(sphingobium.by.HII.GAM)
# Visual check = assumptions met

## Sphingobium abundance-by-ISC model diagnostics
check_model(sphingobium.by.ISC.GAM)
# Visual check = assumptions met
```


ANOVAs

Table 34: ANOVA table for the Sphingobium relative abundance-by-distance GAM. Adjusted R-squared = 0.520, deviance = 53.0 Compartment: F = 74.61, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.528 | 9 | 0.122 | 0.152 |
| s(Distance):CompartmentSoil | 0.000 | 9 | 0.000 | 0.995 |

Table 35: ANOVA table for the Sphingobium relative abundance-by-HII GAM. Adjusted R-squared = 0.523, deviance = 53.4 Compartment: F = 75.13, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.613 | 9 | 0.176 | 0.113 |
| s(Human_Influence_Index):CompartmentSoil | 0.000 | 9 | 0.000 | 0.933 |

Table 36: ANOVA table for the Sphingobium relative abundance-by-ISC GAM. Adjusted R-squared = 0.512, deviance = 51.9 Compartment: F = 73.42, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-----|---------|---|---------|
| s(Mean_ISC):CompartmentRoot | 0 | 9 | 0 | 0.932 |
| s(Mean_ISC):CompartmentSoil | 0 | 9 | 0 | 0.993 |

Desulfosporosinus GAMs

```
## Desulfosporosinus by distance
desulfosporosinus.by.distance.GAM <- gam(
  Desulfosporosinus_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfosporosinus by HII
desulfosporosinus.by.HII.GAM <- gam(
  Desulfosporosinus_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfosporosinus by ISC
desulfosporosinus.by.ISC.GAM <- gam(
  Desulfosporosinus_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Desulfosporosinus abundance-by-distance model diagnostics
check_model(desulfosporosinus.by.distance.GAM)
# Visual check = assumptions met

## Desulfosporosinus abundance-by-HII model diagnostics
check_model(desulfosporosinus.by.HII.GAM)
# Visual check = assumptions met

## Desulfosporosinus abundance-by-ISC model diagnostics
check_model(desulfosporosinus.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 37: ANOVA table for the Desulfosporosinus relative abundance-by-distance GAM. Adjusted R-squared = 0.442, deviance = 46.4 Compartment: F = 50.04, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.001 | 9 | 0.000 | 0.704 |
| s(Distance):CompartmentSoil | 1.641 | 9 | 0.632 | 0.034 |

Table 38: ANOVA table for the Desulfosporosinus relative abundance-by-HII GAM. Adjusted R-squared = 0.477, deviance = 50.7 Compartment: F = 53.37, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.633 |
| s(Human_Influence_Index):CompartmentSoil | 2.957 | 9 | 1.176 | 0.013 |

Table 39: ANOVA table for the Desulfosporosinus relative abundance-by-ISC GAM. Adjusted R-squared = 0.572, deviance = 61.2 Compartment: F = 65.19, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.00 | 9 | 0.000 | 0.581 |
| s(Mean_ISC):CompartmentSoil | 5.46 | 9 | 3.111 | 0.000 |

Desulfocapsa GAMs

```
## Desulfocapsa by distance
desulfocapsa.by.distance.GAM <- gam(
  Desulfocapsa_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfocapsa by HII
desulfocapsa.by.HII.GAM <- gam(
  Desulfocapsa_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Desulfocapsa by ISC
desulfocapsa.by.ISC.GAM <- gam(
  Desulfocapsa_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Desulfocapsa abundance-by-distance model diagnostics
check_model(desulfocapsa.by.distance.GAM)
# Visual check = assumptions met

## Desulfocapsa abundance-by-HII model diagnostics
check_model(desulfocapsa.by.HII.GAM)
# Visual check = assumptions met

## Desulfocapsa abundance-by-ISC model diagnostics
check_model(desulfocapsa.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 40: ANOVA table for the *Desulfocapsa* relative abundance-by-distance GAM. Adjusted R-squared = 0.230, deviance = 25.1 Compartment: F = 14.22, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.00 | 9 | 0.000 | 0.999 |
| s(Distance):CompartmentSoil | 0.92 | 9 | 0.816 | 0.005 |

Table 41: ANOVA table for the *Desulfocapsa* relative abundance-by-HII GAM. Adjusted R-squared = 0.230, deviance = 25.1 Compartment: F = 14.23, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.000 | 9 | 0.000 | 0.971 |
| s(Human_Influence_Index):CompartmentSoil | 0.881 | 9 | 0.823 | 0.005 |

Table 42: ANOVA table for the *Desulfocapsa* relative abundance-by-ISC GAM. Adjusted R-squared = 0.224, deviance = 25.7 Compartment: F = 14.12, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.000 | 9 | 0.000 | 0.999 |
| s(Mean_ISC):CompartmentSoil | 1.858 | 9 | 0.761 | 0.024 |

Phyllobacterium GAMs

```
## Phyllobacterium by distance
phyllobacterium.by.distance.GAM <- gam(
  Phyllobacterium_Relative_Abundance ~ Compartment
  + s(Distance, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Phyllobacterium by HII
phyllobacterium.by.HII.GAM <- gam(
  Phyllobacterium_Relative_Abundance ~ Compartment
  + s(Human_Influence_Index, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)

## Phyllobacterium by ISC
phyllobacterium.by.ISC.GAM <- gam(
  Phyllobacterium_Relative_Abundance ~ Compartment
  + s(Mean_ISC, by = Compartment, bs = "ts", k = 10),
  data = ecosystem.function.bacteria.data,
  method = "REML"
)
```

Check Model Assumptions

```
## Phyllobacterium abundance-by-distance model diagnostics
check_model(phyllobacterium.by.distance.GAM)
# Visual check = assumptions met

## Phyllobacterium abundance-by-HII model diagnostics
check_model(phyllobacterium.by.HII.GAM)
# Visual check = assumptions met

## Phyllobacterium abundance-by-ISC model diagnostics
check_model(phyllobacterium.by.ISC.GAM)
# Visual check = assumptions met
```

ANOVAs

Table 43: ANOVA table for the *Phyllobacterium* relative abundance-by-distance GAM. Adjusted R-squared = 0.358, deviance = 37.0 Compartment: F = 38.94, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Distance):CompartmentRoot | 0.339 | 9 | 0.056 | 0.225 |
| s(Distance):CompartmentSoil | 0.000 | 9 | 0.000 | 0.905 |

Table 44: ANOVA table for the *Phyllobacterium* relative abundance-by-HII GAM. Adjusted R-squared = 0.399, deviance = 41.5 Compartment: F = 41.59, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|--|-------|---------|-------|---------|
| s(Human_Influence_Index):CompartmentRoot | 0.838 | 9 | 0.574 | 0.016 |
| s(Human_Influence_Index):CompartmentSoil | 0.000 | 9 | 0.000 | 0.695 |

Table 45: ANOVA table for the *Phyllobacterium* relative abundance-by-ISC GAM. Adjusted R-squared = 0.392, deviance = 40.8 Compartment: F = 41.15, P < 0.001.

| Term | EDF | Ref. df | F | P-value |
|-----------------------------|-------|---------|-------|---------|
| s(Mean_ISC):CompartmentRoot | 0.834 | 9 | 0.487 | 0.024 |
| s(Mean_ISC):CompartmentSoil | 0.000 | 9 | 0.000 | 0.566 |

Export Data

```
## Ecosystem function bacteria
write_rds(
  ecosystem.function.bacteria.data,
  file = "data/ecosystem_function_bacteria_data.rds"
)
```


R Session Information

Table 46: Packages required for data management and analysis.

| Package | Loaded Version | Date |
|---------------|----------------|------------|
| bayestestR | 0.13.2 | 2024-02-12 |
| broom | 1.0.5 | 2023-06-09 |
| correlation | 0.8.4 | 2023-04-06 |
| datawizard | 0.10.0 | 2024-03-26 |
| dplyr | 1.1.4 | 2023-11-17 |
| easystats | 0.7.1 | 2024-03-28 |
| effectsize | 0.8.7 | 2024-04-01 |
| forcats | 1.0.0 | 2023-01-29 |
| ggplot2 | 3.5.1 | 2024-04-23 |
| insight | 0.19.10 | 2024-03-22 |
| kableExtra | 1.4.0 | 2024-01-24 |
| knitr | 1.46 | 2024-04-06 |
| lubridate | 1.9.3 | 2023-09-27 |
| mgcv | 1.9-1 | 2023-12-21 |
| modelbased | 0.8.7 | 2024-02-15 |
| nlme | 3.1-164 | 2023-11-27 |
| parameters | 0.21.6 | 2024-03-18 |
| performance | 0.11.0 | 2024-03-22 |
| purrr | 1.0.2 | 2023-08-10 |
| readr | 2.1.5 | 2024-01-10 |
| report | 0.5.8 | 2023-12-07 |
| see | 0.8.4 | 2024-04-29 |
| stringr | 1.5.1 | 2023-11-14 |
| tibble | 3.2.1 | 2023-03-20 |
| tidyamplicons | 0.2.2 | 2022-09-10 |
| tidyr | 1.3.1 | 2024-01-24 |
| tidyverse | 2.0.0 | 2023-02-22 |